SEQ ID NO: 1

```
                                        SUMMARIES
                         %
    Result              Query
       No.    Score     Match  Length  DB   ID                       Description
    ------------------------------------------------------------------------------------
         1     6880     100.0    6880   6   AX253450            AX253450 Sequence
         2     6880     100.0    6880   9   HSA012376           AJ012376 Homo sapi
         3     6880     100.0    7260   6   AX253452            AX253452 Sequence
         4     6879     100.0    9497   6   AX059973            AX059973 Sequence
         5     6879     100.0    9497   9   AF165281            AF165281 Homo sapi
         6   6875.8      99.9    9741   6   AX127830            AX127830 Sequence
         7   6875.8      99.9    9741   6   AX139817            AX139817 Sequence
         8   6875.8      99.9    9741   6   AX351038            AX351038 Sequence
         9   6875.8      99.9    9854   6   AX127831            AX127831 Sequence
        10   6875.8      99.9    9854   6   AX139818            AX139818 Sequence
        11   6870.4      99.9    7860   6   AX092594            AX092594 Sequence
        12     6865      99.8    9495   6   AX059978            AX059978 Sequence
        13   6862.4      99.7   10442   6   AX060713            AX060713 Sequence
        14   6862.4      99.7   10442   6   AX060892            AX060892 Sequence
        15   6862.4      99.7   10442   9   AF285167            AF285167 Homo sapi
        16   6860.8      99.7   10474   6   AX060719            AX060719 Sequence
        17   6860.8      99.7   10474   6   AX060721            AX060721 Sequence
        18   6860.8      99.7   10474   6   AX060898            AX060898 Sequence
        19   6860.8      99.7   10474   6   AX060900            AX060900 Sequence
        20   6845.2      99.5    7862   6   AX135712            AX135712 Sequence
        21   6760.2      98.3    9593   6   AX059976            AX059976 Sequence
        22   6711.6      97.6    6786   9   AB055982            AB055982 Homo sapi
        23     5630      81.8    7878  10   MMABC1               X75926 Mus musculu
        24   5455.6      79.3    6801  10   AY208182            AY208182 Rattus no
        25     4302      62.5    7074   5   AF362377            AF362377 Gallus ga
        26   2052.4      29.8    5097   6   BD012346            BD012346 Genes rel
        27   2052.4      29.8    5097   6   BD160225            BD160225 Primer fo
        28   2052.4      29.8    5097   9   AK027864            AK027864 Homo sapi
        29   1911.6      27.8    7298   9   AF001945            AF001945 Homo sapi
```

```
AF165281
LOCUS       AF165281              9497 bp    mRNA    linear   PRI 17-AUG-1999
DEFINITION  Homo sapiens ATP cassette binding transporter 1 (ABC1) mRNA,
            complete cds.
ACCESSION   AF165281
VERSION     AF165281.1  GI:5734100
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9497)
  AUTHORS   Rust,S., Rosier,M., Funke,H., Real,J., Amoura,Z., Piette,J.C.,
            Deleuze,J.F., Brewer,H.B., Duverger,N., Denefle,P  and Assmann,G.
  TITLE     Tangier disease is caused by mutations in the gene encoding
            ATP-binding cassette transporter 1
  JOURNAL   Nat. Genet. 22 (4), 352-355 (1999)
  MEDLINE   99364413
   PUBMED   10431238
REFERENCE   2  (bases 1 to 9497)
  AUTHORS   Rust,S., Rosier,M., Funke,H., Real,J., Amoura,Z., Piette,J.C.,
            Deleuze,J.F., Brewer,H.B., Duverger,N., Denefle,P. and Assmann,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-1999) Genomics, Rhone-Poulenc Rorer, 2 rue Gaston
            Cr#mieux, Evry 91006, France
FEATURES             Location/Qualifiers
     source          1..9497
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                        /db_xref="taxon:9606"
                        /chromosome="9"
                        /map="9q31"
        gene            1. .9497
                        /gene="ABC1"
        CDS             121. .6726
                        /gene="ABC1"
                        /note="ABC transporter; ABC1 protein"
                        /codon_start=1
                        /product="ATP cassette binding transporter 1"
                        /protein_id="AAD49849.1"
                        /db_xref="GI:5734101"
                        /translation="MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVAR
                        LFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYHNL
                        SLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLPRE
                        KLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMRSW
                        SDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGG
                        NGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATRQ
                        VMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQ
                        LDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLNKL
                        EPIATEVVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERTNK
                        IKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPYPC
                        YVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILWFS
                        WFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLFSR
                        ANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASLLSPVAFGFGCEYFALFEEQ
                        GIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTWYIEAVFPGQYGIPRPWYF
                        PCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVSIQNLVKVYRDGMKVAV
                        DGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQN
                        LGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKTSQ
                        LSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILSTHH
                        MDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSSST
                        VSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHELTY
                        VLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETSDG
                        TLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVKGW
                        KLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQ
                        PWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEWTT
                        APVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTADI
                        LQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVNDA
                        TKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNA
                        ILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMSFVPA
                        SFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKS
                        YVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATF
                        VLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPL
                        SWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRI
                        LDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLT
                        GDTTVTRGDAFLNRNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPE
                        KEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
                        PKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNR
                        FGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPKEKHRNMLQYQLPSSLSSLARIFS
                        ILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSF
                        LQDEKVKESYV"
BASE COUNT      2600 a   2115 c   2217 g   2564 t        1 others
ORIGIN

   Query Match              100.0%;  Score 6879;  DB 9;  Length 9497;
   Best Local Similarity    100.0%;  Pred. No. 0;
   Matches 6879;  Conservative   0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCTTCCTGATCCTGATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCTTCCTGATCCTGATC 60

Qy     61 TCTGTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATTTTCCAAATAAAGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTGTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATTTTCCAAATAAAGCC 120

Qy    121 ATGCCCTCTGCAGGAACACTTCCTTGGGTTCAGGGGATTATCTGTAATGCCAACAACCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATGCCCTCTGCAGGAACACTTCCTTGGGTTCAGGGGATTATCTGTAATGCCAACAACCCC 180

Qy    181 TGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGGAGTTGTTGGAAACTTTAACAAATCC 240
```

```
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGGAGTTGTTGGAAACTTTAACAAATCC 240

Qy      241 ATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACC 300
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACC 300

Qy      301 AGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCA 360
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCA 360

Qy      361 AACTTGAAGCTTCAAGATTTCCTGGTGGACAATGAAACCTTCTCTGGGTTCCTGTATCAC 420
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AACTTGAAGCTTCAAGATTTCCTGGTGGACAATGAAACCTTCTCTGGGTTCCTGTATCAC 420

Qy      421 AACCTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTGATGTCATTCTCCAC 480
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AACCTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTGATGTCATTCTCCAC 480

Qy      481 AAGGTATTTTTGCAAGGCTACCAGTTACATTTGACAAGTCTGTGCAATGGATCAAAATCA 540
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AAGGTATTTTTGCAAGGCTACCAGTTACATTTGACAAGTCTGTGCAATGGATCAAAATCA 540

Qy      541 GAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTGGCCTACCAAGGGAG 600
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTGGCCTACCAAGGGAG 600

Qy      601 AAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCCAATCCTG 660
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCCAATCCTG 660

Qy      661 AGAACACTAAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCCGAAGCCACAAAAACA 720
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AGAACACTAAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCCGAAGCCACAAAAACA 720

Qy      721 TTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTCAGCATGAGAAGCTGGAGTGAC 780
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTCAGCATGAGAAGCTGGAGTGAC 780

Qy      781 ATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCTCCTCCACCCAAATC 840
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 ATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCTCCTCCACCCAAATC 840

Qy      841 TACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGGGGCTGAAGATCAAG 900
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGGGGCTGAAGATCAAG 900

Qy      901 TCTCTCAACTGGTATGAGGACAACAACTACAAAGCCCTCTTTGGAGGCAATGGCACTGAG 960
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TCTCTCAACTGGTATGAGGACAACAACTACAAAGCCCTCTTTGGAGGCAATGGCACTGAG 960

Qy      961 GAAGATGCTGAAACCTTCTATGACAACTCTACAACTCCTTACTGCAATGATTTGATGAAG 1020
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GAAGATGCTGAAACCTTCTATGACAACTCTACAACTCCTTACTGCAATGATTTGATGAAG 1020

Qy     1021 AATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGAAGCCGCTGCTCGTT 1080
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 AATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGAAGCCGCTGCTCGTT 1080

Qy     1081 GGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCATGGCTGAGGTGAAC 1140
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCATGGCTGAGGTⸯAAC 1140

Qy     1141 AAGACCTTCCAGGAACTGGCTGTGTTCCATGATCTGGAAGGCATGTGGGAGGAACTCAGC 1200
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AAGACCTTCCAGGAACTGGCTGTGTTCCATGATCTGGAAGGCATGTGGGAGGAACTCAGC 1200

Qy     1201 CCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTGTCCGGATGCTGTTG 1260
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTGⸯCCGGATGCTGTTG 1260
```

```
Qy    1261 GACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCC 1320

Qy    1321 CAAGACATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CAAGACATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCT 1380

Qy    1381 GTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCCGGACCATATCTCGC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCCGGACCATATCTCGC 1440

Qy    1441 TTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAGAAGTCTGGCTCATC 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAGAAGTCTGGCTCATC 1500

Qy    1501 AACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTCACTGGA 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTCACTGGA 1560

Qy    1561 ATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGATCCGAATGGACATT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 ATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGATCCGAATGGACATT 1620

Qy    1621 GACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACCCTGGTCCTCGAGCT 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACCCTGGTCCTCGAGCT 1680

Qy    1681 GACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGCTTCGCCTACTTGCAGGATGTGGTG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGCTTCGCCTACTTGCAGGATGTGGTG 1740

Qy    1741 GAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAACTGGTGTCTATATGCAA 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAACTGGTGTCTATATGCAA 1800

Qy    1801 CAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGATGAGCCGGTCAATG 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 CAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGATGAGCCGGTCAATG 1860

Qy    1861 CCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCATCAAGGGCATCGTG 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 CCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCATCAAGGGCATCGTG 1920

Qy    1921 TATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCCTGGACAACAGCATC 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 TATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCCTGGACAACAGCATC 1980

Qy    1981 CTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCCTCTTCTTGTGAGCGCTGGCCTGCTA 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 CTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCCTCTTCTTGTGAGCGCTGGCCTGCTA 2040

Qy    2041 GTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCAGCGTGGTGTTTGTC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 GTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCAGCGTGGTGTTTGTC 2100

Qy    2101 TTCCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGATTAGCACACTCTTC 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 TTCCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGATTAGCACACTCTTC 2160

Qy    2161 TCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCACGCTGTACCTGCCC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 TCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCACGCTGTACCTGCCC 2220

Qy    2221 TACGTCCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCAAGATCTTCGCTAGC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 TACGTCCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCAAGATCTTCGCTAGC 2280

Qy    2281 CTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCCTTTTTGAGGAGCAG 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCCTTTTTGAGGAGCAG 2340
```

```
Qy    2341 GGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGGAAGATGGCTTCAAT 2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGGAAGATGGCTTCAAT 2400

Qy    2401 CTCACCACTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATGGGGTGATGACCTGG 2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 CTCACCACTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATGGGGTGATGACCTGG 2460

Qy    2461 TACATTGAGGCTGTCTTTCCAGGCCAGTACGGAATTCCCAGGCCCTGGTATTTTCCTTGC 2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 TACATTGAGGCTGTCTTTCCAGGCCAGTACGGAATTCCCAGGCCCTGGTATTTTCCTTGC 2520

Qy    2521 ACCAAGTCCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACCCTGGTTCCAACCAG 2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ACCAAGTCCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACCCTGGTTCCAACCAG 2580

Qy    2581 AAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGAAGCTGGGCGTGTCC 2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 AAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGAAGCTGGGCGTGTCC 2640

Qy    2641 ATTCAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTCGATGGCCTGGCA 2700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 ATTCAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTCGATGGCCTGGCA 2700

Qy    2701 CTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATGGAGCGGGGAAGACG 2760
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATGGAGCGGGGAAGACG 2760

Qy    2761 ACCACCATGTCAATCCTGACCGGGTTGTTCCCCCCGACCTCGGGCACCGCCTACATCCTG 2820
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 ACCACCATGTCAATCCTGACCGGGTTGTTCCCCCCGACCTCGGGCACCGCCTACATCCTG 2820

Qy    2821 GGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTCTGTCCCCAG 2880
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTCTGTCCCCAG 2880

Qy    2881 CATAACGTGCTGTTTGACATGCTGACTGTCGAAGAACACATCTGGTTCTATGCCCGCTTG 2940
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 CATAACGTGCTGTTTGACATGCTGACTGTCGAAGAACACATCTGGTTCTATGCCCGCTTG 2940

Qy    2941 AAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGGCCCTGGATGTTGGT 3000
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 AAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGGCCCTGGATGTTGGT 3000

Qy    3001 TTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCAGGTGGAATGCAGAGAAAG 3060
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 TTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCAGGTGGAATGCAGAGAAAG 3060

Qy    3061 CTATCTGTGGCCTTGGCCTTTGTCGGGGGATCTAAGGTTGTCATTCTGGATGAACCCACA 3120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 CTATCTGTGGCCTTGGCCTTTGTCGGGGGATCTAAGGTTGTCATTCTGGATGAACCCACA 3120

Qy    3121 GCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTGCTGCTGAAATACCGACAA 3180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 GCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTGCTGCTGAAATACCGACAA 3180

Qy    3181 GGCCGCACCATTATTCTCTCTACACACCACATGGATGAAGCGGACGTCCTGGGGGACAGG 3240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3181 GGCCGCACCATTATTCTCTCTACACACCACATGGATGAAGCGGACGTCCTGGGGGACAGG 3240

Qy    3241 ATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTCTGAAGAAC 3300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 ATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTCTGAAGAAC 3300

Qy    3301 CAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGT 3360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3301 CAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGT 3360

Qy    3361 TCCTGCAGAAACAGTAGTAGCACTGTGTCATACCTGAAAAAGGAGGACAGTGTTTCTCAG 3420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   3361 TCCTGCAGAAACAGTAGTAGCACTGTGTCATACCTGAAAAAGGAGGACAGTGTTTCTCAG 3420

Qy   3421 AGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGATGTC 3480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 AGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGATGTC 3480

Qy   3481 TCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCGGCTGGTGGAAGACATA 3540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 TCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCGGCTGGTGGAAGACATA 3540

Qy   3541 GGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTTGTGGAA 3600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTTGTGGAA 3600

Qy   3601 CTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCA 3660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 CTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCA 3660

Qy   3661 GAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAG 3720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAG 3720

Qy   3721 ACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCGGGGACAAGCAGAGC 3780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCGGGGACAAGCAGAGC 3780

Qy   3781 TGTCTTCGCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATTCTGACATAGACCCA 3840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 TGTCTTCGCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATTCTGACATAGACCCA 3840

Qy   3841 GAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGTCCTACCAGGTGAAA 3900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGTCCTACCAGGTGAAA 3900

Qy   3901 GGCTGGAAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGACTGCTAATTGCC 3960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GGCTGGAAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGACTGCTAATTGCC 3960

Qy   3961 AGACGGAGTCGGAAAGGATTTTTTGCTCAGATTGTCTTGCCAGCTGTGTTTGTCTGCATT 4020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 AGACGGAGTCGGAAAGGATTTTTTGCTCAGATTGTCTTGCCAGCTGTGTTTGTCTGCATT 4020

Qy   4021 GCCCTTGTGTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCCAGCCTGGAACTTCAG 4080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GCCCTTGTGTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCCAGCCTGGAACTTCAG 4080

Qy   4081 CCCTGGATGTACAACGAACAGTACACATTTGTCAGCAATGATGCTCCTGAGGACACGGGA 4140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 CCCTGGATGTACAACGAACAGTACACATTTGTCAGCAATGATGCTCCTGAGGACACGGGA 4140

Qy   4141 ACCCTGGAACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAA 4200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ACCCTGGAACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAA 4200

Qy   4201 GGAAACCCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGTGGACCACTGCCCCA 4260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGAAACCCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGTGGACCACTGCCCCA 4260

Qy   4261 GTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACTGGACAATGCAGAACCCTTCA 4320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 GTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACTGGACAATGCAGAACCCTTCA 4320

Qy   4321 CCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTGTGTGTCCCCCAGGG 4380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTGTGTGTCCCCCAGGG 4380

Qy   4381 GCAGGGGGGCTGCCTCCTCCACAAAGAAAACAAAACACTGCAGATATCCTTCAGGACCTG 4440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 GCAGGGGGGCTGCCTCCTCCACAAAGAAAACAAAACACTGCAGATATCCTTCAGGACCTG 4440

Qy   4441 ACAGGAAGAAACATTTCGGATTATCTGGTGAAGACGTATGTGCAGATCATAGCCAAAAGC 4500

```
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ACAGGAAGAAACATTTCGGATTATCTGGTGAAGACGTATGTGCAGATCATAGCCAAAAGC 4500

Qy   4501 TTAAAGAACAAGATCTGGGTGAATGAGTTTAGGTATGGCGGCTTTTCCCTGGGTGTCAGT 4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TTAAAGAACAAGATCTGGGTGAATGAGTTTAGGTATGGCGGCTTTTCCCTGGGTGTCAGT 4560

Qy   4561 AATACTCAAGCACTTCCTCCGAGTCAAGAAGTTAATGATGCCACCAAACAAATGAAGAAA 4620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 AATACTCAAGCACTTCCTCCGAGTCAAGAAGTTAATGATGCCACCAAACAAATGAAGAAA 4620

Qy   4621 CACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACAGCTTGGGAAGATTT 4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 CACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACAGCTTGGGAAGATTT 4680

Qy   4681 ATGACAGGACTGGACACCAGAAATAATGTCAAGGTGTGGTTCAATAACAAGGGCTGGCAT 4740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 ATGACAGGACTGGACACCAGAAATAATGTCAAGGTGTGGTTCAATAACAAGGGCTGGCAT 4740

Qy   4741 GCAATCAGCTCTTTCCTGAATGTCATCAACAATGCCATTCTCCGGGCCAACCTGCAAAAG 4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GCAATCAGCTCTTTCCTGAATGTCATCAACAATGCCATTCTCCGGGCCAACCTGCAAAAG 4800

Qy   4801 GGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCCTGAATCTCACCAAG 4860
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   4801 GGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCCTGAATCTCACCAAG 4860

Qy   4861 CAGCAGCTCTCAGAGGTGGCTCCGATGACCACATCAGTGGATGTCCTTGTGTCCATCTGT 4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CAGCAGCTCTCAGAGGTGGCTCCGATGACCACATCAGTGGATGTCCTTGTGTCCATCTGT 4920

Qy   4921 GTCATCTTTGCAATGTCCTTCGTCCCAGCCAGCTTTGTCGTATTCCTGATCCAGGAGCGG 4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 GTCATCTTTGCAATGTCCTTCGTCCCAGCCAGCTTTGTCGTATTCCTGATCCAGGAGCGG 4980

Qy   4981 GTCAGCAAAGCAAAACACCTGCAGTTCATCAGTGGAGTGAAGCCTGTCATCTACTGGCTC 5040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 GTCAGCAAAGCAAAACACCTGCAGTTCATCAGTGGAGTGAAGCCTGTCATCTACTGGCTC 5040

Qy   5041 TCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACACTGGTCATTATCATC 5100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACACTGGTCATTATCATC 5100

Qy   5101 TTCATCTGCTTCCAGCAGAAGTCCTATGTGTCCTCCACCAATCTGCCTGTGCTAGCCCTT 5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 TTCATCTGCTTCCAGCAGAAGTCCTATGTGTCCTCCACCAATCTGCCTGTGCTAGCCCTT 5160

Qy   5161 CTACTTTTGCTGTATGGGTGGTCAATCACACCTCTCATGTACCCAGCCTCCTTTGTGTTC 5220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 CTACTTTTGCTGTATGGGTGGTCAATCACACCTCTCATGTACCCAGCCTCCTTTGTGTTC 5220

Qy   5221 AAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCTTCATTGGCATTAAT 5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 AAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCTTCATTGGCATTAAT 5280

Qy   5281 GGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTCACCGACAATAAGCTGAATAATATCAAT 5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTCACCGACAATAAGCTGAATAATATCAAT 5340

Qy   5341 GATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTTGCCTGGGACGAGGGCTCATC 5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTTGCCTGGGACGAGGGCTCATC 5400

Qy   5401 GACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTGGGGAGAATCGCTTT 5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 GACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTGGGGAGAATCGCTTT 5460

Qy   5461 GTGTCACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCATGGCCGTGGAAGGG 5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 GTGTCACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCATGGCCGTGGAAGGG 5520
```

```
Qy    5521 GTGGTGTTCTTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCATCAGGCCCAGACCT 5580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5521 GTGGTGTTCTTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCATCAGGCCCAGACCT 5580

Qy    5581 GTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGATGTGAGGCGGGAAAGACAG 5640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5581 GTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGATGTGAGGCGGGAAAGACAG 5640

Qy    5641 AGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGACGAAGATA 5700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5641 AGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGACGAAGATA 5700

Qy    5701 TATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCATTCCTCCTGGTGAG 5760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5701 TATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCATTCCTCCTGGTGAG 5760

Qy    5761 TGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGAAAATCATCAACTTTCAAGATGTTAACA 5820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5761 TGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGAAAATCATCAACTTTCAAGATGTTAACA 5820

Qy    5821 GGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAGAAATAGTATCTTATCAAAC 5880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5821 GGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAGAAATAGTATCTTATCAAAC 5880

Qy    5881 ATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATGCCATCACAGAGCTG 5940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5881 ATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATGCCATCACAGAGCTG 5940

Qy    5941 TTGACTGGGAGAGAACACGTGGAGTTCTTTGCCCTTTTGAGAGGAGTCCCAGAGAAAGAA 6000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5941 TTGACTGGGAGAGAACACGTGGAGTTCTTTGCCCTTTTGAGAGGAGTCCCAGAGAAAGAA 6000

Qy    6001 GTTGGCAAGGTTGGTGAGTGGGCGATTCGGAAACTGGGCCTCGTGAAGTATGGAGAAAAA 6060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6001 GTTGGCAAGGTTGGTGAGTGGGCGATTCGGAAACTGGGCCTCGTGAAGTATGGAGAAAAA 6060

Qy    6061 TATGCTGGTAACTATAGTGGAGGCAACAAACGCAAGCTCTCTACAGCCATGGCTTTGATC 6120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6061 TATGCTGGTAACTATAGTGGAGGCAACAAACGCAAGCTCTCTACAGCCATGGCTTTGATC 6120

Qy    6121 GGCGGGCCTCCTGTGGTGTTTCTGGATGAACCCACCACAGGCATGGATCCCAAAGCCCGG 6180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6121 GGCGGGCCTCCTGTGGTGTTTCTGGATGAACCCACCACAGGCATGGATCCCAAAGCCCGG 6180

Qy    6181 CGGTTCTTGTGGAATTGTGCCCTAAGTGTTGTCAAGGAGGGGAGATCAGTAGTGCTTACA 6240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6181 CGGTTCTTGTGGAATTGTGCCCTAAGTGTTGTCAAGGAGGGGAGATCAGTAGTGCTTACA 6240

Qy    6241 TCTCATAGTATGGAAGAATGTGAAGCTCTTTGCACTAGGATGGCAATCATGGTCAATGGA 6300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6241 TCTCATAGTATGGAAGAATGTGAAGCTCTTTGCACTAGGATGGCAATCATGGTCAATGGA 6300

Qy    6301 AGGTTCAGGTGCCTTGGCAGTGTCCAGCATCTAAAAAAATAGGTTTGGAGATGGTTATACA 6360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6301 AGGTTCAGGTGCCTTGGCAGTGTCCAGCATCTAAAAAAATAGGTTTGGAGATGGTTATACA 6360

Qy    6361 ATAGTTGTACGAATAGCAGGGTCCAACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGGA 6420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6361 ATAGTTGTACGAATAGCAGGGTCCAACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGGA 6420

Qy    6421 CTTGCATTTCCTGGAAGTGTTCCAAAAGAGAAACACCGGAACATGCTACAATACCAGCTT 6480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6421 CTTGCATTTCCTGGAAGTGTTCCAAAAGAGAAACACCGGAACATGCTACAATACCAGCTT 6480

Qy    6481 CCATCTTCATTATCTTCTCTGGCCAGGATATTCAGCATCCTCTCCCAGAGCAAAAAGCGA 6540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6481 CCATCTTCATTATCTTCTCTGGCCAGGATATTCAGCATCCTCTCCCAGAGCAAAAAGCGA 6540

Qy    6541 CTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAGTATTTGTGAACTTT 6600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6541 CTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAGTATTTGTGAACTTT 6600
```

```
Qy   6601 GCCAAGGACCAAAGTGATGATGACCACTTAAAAGACCTCTCATTACACAAAAACCAGACA 6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCCAAGGACCAAAGTGATGATGACCACTTAAAAGACCTCTCATTACACAAAAACCAGACA 6660

Qy   6661 GTAGTGGACGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAGTGAAAGAAAGCTAT 6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 GTAGTGGACGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAGTGAAAGAAAGCTAT 6720

Qy   6721 GTATGAAGAATCCTGTTCATACGGGGTGGCTGAAAGTAAAGAGGGACTAGACTTTCCTTT 6780
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   6721 GTATGAAGAATCCTGTTCATACGGGGTGGCTGAAAGTAAAGAGGNACTAGACTTTCCTTT 6780

Qy   6781 GCACCATGTGAAGTGTTGTGGAGAAAAGAGCCAGAAGTTGATGTGGGAAGAAGTAAACTG 6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 GCACCATGTGAAGTGTTGTGGAGAAAAGAGCCAGAAGTTGATGTGGGAAGAAGTAAACTG 6840

Qy   6841 GATACTGTACTGATACTATTCAATGCAATGCAATTCAATG 6880
          ||||||||||||||||||||||||||||||||||||||||
Db   6841 GATACTGTACTGATACTATTCAATGCAATGCAATTCAATG 6880
```

<div align="center">SUMMARIES</div>

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6880	100.0	6880	21	AAZ94734	Human ATP binding
2	6880	100.0	6880	22	AAI70314	Human ATP binding
3	6880	100.0	7260	22	AAD21326	Human ATP binding
4	6880	100.0	7260	22	AAI70315	Human ATP binding
5	6879	100.0	9497	24	ABV78020	Hypoxia-regulated
6	6876.8	100.0	6880	22	AAD21325	Human ATP binding
7	6875.8	99.9	9741	22	AAS06120	Human ABC1 DNA seq
8	6875.8	99.9	9741	24	AAD37273	Human ABC1 full-le
9	6875.8	99.9	9854	22	AAS06121	Human ABC1 DNA seq
10	6875.8	99.9	9870	24	ABN99301	Polymorphic human
11	6875.8	99.9	9870	24	ABN99302	Polymorphic human
12	6875.8	99.9	9870	24	ABN99303	Polymorphic human
13	6875.8	99.9	9870	24	ABN99324	Polymorphic human
14	6875.8	99.9	9870	24	ABN99328	Polymorphic human
15	6875.8	99.9	9870	24	ABN99329	Polymorphic human
16	6875.8	99.9	9870	24	ABN99330	Polymorphic human
17	6875.8	99.9	9870	24	ABN99331	Polymorphic human
18	6875.8	99.9	9870	24	ABN99332	Polymorphic human
19	6875.8	99.9	9870	24	ABN99333	Polymorphic human
20	6875.8	99.9	9870	24	ABN99334	Polymorphic human
21	6875.8	99.9	9870	24	ABN99335	Polymorphic human
22	6875.2	99.9	7281	22	AAK51683	Human polynucleoti
23	6874.2	99.9	9741	24	ABL58146	Human ABCA1 transp
24	6874.2	99.9	9870	24	ABN99304	Polymorphic human
25	6874.2	99.9	9870	24	ABN99305	Polymorphic human
26	6874.2	99.9	9870	24	ABN99306	Polymorphic human
27	6874.2	99.9	9870	24	ABN99307	Polymorphic human
28	6874.2	99.9	9870	24	ABN99309	Polymorphic human
29	6874.2	99.9	9870	24	ABN99310	Polymorphic human
30	6874.2	99.9	9870	24	ABN99311	Polymorphic human
31	6874.2	99.9	9870	24	ABN99312	Polymorphic human
32	6874.2	99.9	9870	24	ABN99313	Polymorphic human
33	6874.2	99.9	9870	24	ABN99314	Polymorphic human
34	6874.2	99.9	9870	24	ABN99315	Polymorphic human
35	6874.2	99.9	9870	24	ABN99316	Polymorphic human
36	6874.2	99.9	9870	24	ABN99317	Polymorphic human
37	6874.2	99.9	9870	24	ABN99318	Polymorphic human
38	6874.2	99.9	9870	24	ABN99319	Polymorphic human
39	6874.2	99.9	9870	24	ABN99320	Polymorphic human
40	6874.2	99.9	9870	24	ABN99321	Polymorphic human
41	6874.2	99.9	9870	24	ABN99322	Polymorphic human
42	6874.2	99.9	9870	24	ABN99323	Polymorphic human
43	6874.2	99.9	9870	24	ABN99325	Polymorphic human
44	6874.2	99.9	9870	24	ABN99326	Polymorphic human

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	430.8	6.3	5894	3	US-08-665-259-24	Sequence 24, Appl
2	430.8	6.3	5894	3	US-08-762-500-24	Sequence 24, Appl
3	430.8	6.3	6525	3	US-08-762-500-74	Sequence 74, Appl
4	99.4	1.4	1008	4	US-09-252-991A-4266	Sequence 4266, Ap
c 5	98.8	1.4	999	4	US-09-252-991A-4588	Sequence 4588, Ap
6	89.4	1.3	1614	4	US-09-252-991A-4198	Sequence 4198, Ap
7	85.8	1.2	1824	4	US-09-252-991A-1780	Sequence 1780, Ap
8	83.8	1.2	1176	4	US-09-252-991A-3704	Sequence 3704, Ap
c 9	83.8	1.2	2253	4	US-09-252-991A-3810	Sequence 3810, Ap
10	81	1.2	36181	4	US-08-311-731A-120	Sequence 120, App
11	78.4	1.1	1548	4	US-09-252-991A-3653	Sequence 3653, Ap
c 12	77.4	1.1	762	4	US-09-252-991A-1938	Sequence 1938, Ap
c 13	76.6	1.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
14	75.2	1.1	969	4	US-09-107-532A-1008	Sequence 1008, Ap
15	75	1.1	1770	4	US-09-328-352-3466	Sequence 3466, Ap

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3600.4	52.3	4783	11	AK051920	AK051920 Mus muscu
2	1024.8	14.9	3605	11	AK052916	AK052916 Mus muscu
3	692.4	10.1	809	12	BI754756	BI754756 603025477
4	636.6	9.3	854	12	BI854140	BI854140 603381449
5	624.4	9.1	754	14	CD351847	CD351847 UI-M-GI0-
c 6	591.2	8.6	878	12	BI182779	BI182779 UNL-P-FN-
7	587.8	8.5	673	14	CD351490	CD351490 UI-M-GI0-
c 8	550.4	8.0	661	14	CB481810	CB481810 jns83_A02
9	547.6	8.0	773	14	CB526974	CB526974 UI-M-FY0-
10	542	7.9	837	10	BF160011	BF160011 601768192
11	536.8	7.8	652	10	BB468374	BB468374 BB468374
12	535.2	7.8	702	12	BI658600	BI658600 603284335
13	532.4	7.7	724	12	BG920223	BG920223 602822304
c 14	528.4	7.7	639	14	CB477340	CB477340 jns20_C11
15	525	7.6	720	12	BI248317	BI248317 602966918
16	516	7.5	620	14	CA889361	CA889361 B0152D10-
17	510.4	7.4	512	13	BX279856	BX279856 BX279856
c 18	505	7.3	614	14	CB478239	CB478239 jns32_B02
19	486.4	7.1	605	14	CB546553	CB546553 AMGNNUC:N
20	479.8	7.0	579	14	CA884367	CA884367 B0109G01-
21	477	6.9	1004	11	BC034824	BC034824 Homo sapi
22	473	6.9	562	10	BE665489	BE665489 154542 MA
23	459.4	6.7	515	10	BF094524	BF094524 MR0-UT004
24	456.8	6.6	605	14	CB581583	CB581583 AMGNNUC:N
25	447.6	6.5	554	10	BF076322	BF076322 225856 MA
26	438.4	6.4	512	14	CB717058	CB717058 AMGNNUC:N
27	435.6	6.3	884	13	BU906084	BU906084 AGENCOURT

```
RESULT 1
AK051920
LOCUS           AK051920                  4783 bp    mRNA    linear    HTC 05-DEC-2002
DEFINITION      Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length
                enriched library, clone:D230019D04 product:ATP-binding cassette,
                sub-family A (ABC1), member 1, full insert sequence.
ACCESSION       AK051920
VERSION         AK051920.1  GI:26342297
KEYWORDS        HTC; CAP trapper.
SOURCE          Mus musculus (house mouse)
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
    TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
    TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
    TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
    TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
    TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 4783)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
```

Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1. .4783
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:D230019D04"
                     /db_xref="taxon:10090"
                     /clone="D230019D04"
                     /tissue_type="eyeball"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="12 days embryo"
     CDS   322. .>4782
                     /note="unnamed protein product; ATP-binding cassette,
                     sub-family A (ABC1), member 1 (MGD|MGI:99607,
                     GB|NM_013454, evidence: BLASTN, 99%, match=4545)
                     putative"
                     /codon_start=1
                     /protein_id="BAC34811.1"
                     /db_xref="GI:26342298"
                     /translation="MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL
                     SYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIV
                     SRLFSDAQRLLLYSQRDTSIKDMHKVLRMLRQIKHPNSNLKLQDFLVDNETFSGFLQH
                     NLSLPRSTVDSLLQANVGLQKVFLQGYQLHLASLCNGSKLEEIIQLGDAEVSALCGLP
                     RKKLDAAERVLRYNMDILKPVVTKLNSTSHLPTQHLAEATTVLLDSLGGLAQELFSTK
                     SWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALF
                     GGNNTEEDVDTFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPAT
                     RQVMAEVNKTFQELAVFHDLEGMWEELSPQIWTFMENSQEMDLVRTLLDSRGNDQFWE
                     QKLDGLDWTAQDIMAFLAKNPEDAQSPNGSVYTWREAFNETNQAIQTISRFMECVNLN
                     KLEPIPTEVRLINKSMELLDERKFWAGIVFTGITPDSVELPHHVKYKIRMDIDNVERT
                     NKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGSEKKTGVYVQQMPY
                     PCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKSIVYEKEARLKETMRIMGLDNGILW
                     FSWFVSSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAMVTILQCFLISTLF
                     SRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFSIKIFASLLSPVAFGFGCEYFALFE
                     EQGIGVQWDNLFESPVEEDGFNLTTAVSMMLFDTFLYGVMTWYIEAVFPGQYGIPRPW
                     YFPCTKSYWFGEEIDEKSHPGSSQKGVSEICMEEEPTHLRLGVSIQNLVKVYRDGMKV
                     AVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSSIR
                     QNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPPSKLKSKT
                     SQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILST
                     HHMDEADILGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSS
                     STVSCLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
                     TYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETS
                     DGTLPARRNRRAFGDKQSCLHPFTEDDAVDPNDSDIDPESRETDLLSGMDGKGSYQLK
                     GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLE
                     LQPWMYNEQYTFVSNDAPEDMGTQELLNALTKDPGFGTRCMEGNPIPDTPCLAGEEDW
                     TISPVPQSIVDLFQNGNWTMKNPSPACQCSSDKIKKMLPVCPPGAGGLPPP"
BASE COUNT   1119 a   1236 c   1310 g   1118 t
ORIGIN

   Query Match   52.3%; Score 3600.4; DB 11; Length 4783;
   Best Local Similarity 88.6%; Pred. No. 0;
   Matches 3901; Conservative   0; Mismatches 501; Indels   0; Gaps   0;

Qy   1 CAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCTTCCTGATCCTGATC 60

```
          ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db    382 CAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTCTTTATCTTCCTGATCCTGATC 441

Qy     61 TCTGTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATTTTCCAAATAAAGCC 120
          |||||  ||  |||||||||||||||  |||||||||||  |||||  |||| ||||||||||
Db    442 TCTGTACGCCTGAGCTACCCACCCTACGAACAACATGAGTGCCACTTTCCGAATAAAGCC 501

Qy    121 ATGCCCTCTGCAGGAACACTTCCTTGGGTTCAGGGGATTATCTGTAATGCCAACAACCCC 180
          ||||| ||||||||||||| ||  ||  |||  |||||||||||||||||||||||||||||
Db    502 ATGCCGTCTGCAGGAACCCTCCCCTGGGTACAGGGGATTATCTGTAATGCCAACAACCCC 561

Qy    181 TGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGGAGTTGTTGGAAACTTTAACAAATCC 240
          ||  |||||||| ||  |||||  ||  ||||||||| |||||||||||||||||||| |||
Db    562 TGCTTCCGTTATCCAACTCCCGGCGAGGCTCCCGGTGTTGTTGGAAACTTTAACAAGTCC 621

Qy    241 ATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACC 300
          ||  |||  ||||||||||||||| ||||| |||||||| ||  ||||||||| | ||| |||
Db    622 ATCGTGTCTCGCCTGTTCTCAGACGCTCAGAGGCTTCTTCTGTACAGCCAAAGAGATACC 681

Qy    301 AGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCA 360
          |||||  |||||||||| ||| || |||| ||||||| ||||  |||||||||| |  ||| ||||
Db    682 AGCATTAAGGACATGCACAAGGTCCTGAGAATGTTACGGCAGATCAAGCATCCCAACTCA 741

Qy    361 AACTTGAAGCTTCAAGATTTCCTGGTGGACAATGAAACCTTCTCTGGGTTCCTGTATCAC 420
          || ||||||||| |||||| ||||||||||||||||||||| ||||||| ||||||| | |||
Db    742 AATTTGAAGCTCCAGGATTTTCTGGTGGACAATGAAACATTCTCTGGATTCCTGCAGCAC 801

Qy    421 AACCTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTGATGTCATTCTCCAC 480
          ||  |  || ||  |||| ||||||||||| ||||  ||||  |||| |||| |||||
Db    802 AATTTGTCCCTTCCAAGATCTACTGTGGACAGCCTGCTGCAGGCGAATGTTGGTCTCCAG 861

Qy    481 AAGGTATTTTTGCAAGGCTACCAGTTACATTTGACAAGTCTGTGCAATGGATCAAAATCA 540
          |||||||||||||||||||||||| ||||||||  || ||||||| || |||||||||| |
Db    862 AAGGTATTTTTGCAAGGCTACCAATTACATTTGGCCAGTCTGTGTAACGGATCAAAATTA 921

Qy    541 GAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTGGCCTACCAAGGGAG 600
          ||||| || ||||| |||||||| || |||||||| ||  ||||| ||||| |||| || ||
Db    922 GAAGAAATTATTCAGCTTGGTGATGCGGAAGTTTCTGCCCTCTGTGGTCTACCGAGGAAG 981

Qy    601 AAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCCAATCCTG 660
          || || | |||||| ||| ||||||| || |||||||||||||||||||||||| |  ||
Db    982 AAGCTCGATGCAGCCGAGAGAGTACTGCGCTACAACATGGACATCCTGAAGCCAGTTGTG 1041

Qy    661 AGAACACTAAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCCGAAGCCACAAAAACA 720
          | || |||||| || ||||||||  |||||| | ||| | ||||| ||||||||| | |
Db   1042 ACAAAACTAAATTCCACATCTCATCTCCCGACCCAGCATCTGGCTGAAGCCACCACAGTG 1101

Qy    721 TTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTCAGCATGAGAAGCTGGAGTGAC 780
          |||||  | || | ||| || ||||||||| |||||||||||||| |  ||||||| |||
Db   1102 TTGCTTGACAGCTTGGGGGGGCCTGGCCCAAGAGCTGTTCAGCACAAAGAGCTGGAGCGAC 1161

Qy    781 ATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCTCCTCCACCCAAATC 840
          ||||| |||||||||||||||||||||||| ||||||||||||||||||||||||| |||
Db   1162 ATGCGGCAGGAGGTGATGTTTCTGACCAACGTGAACAGCTCCAGCTCCTCCACCCAGATC 1221

Qy    841 TACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGGGGCTGAAGATCAAG 900
          ||||||| |||||| ||  || || || || | ||||  |||| ||||| ||||||||||
Db   1222 TACCAGGCAGTGTCCCGCATCGTGTGTGGTCACCCAGAGGGTGGGGGCCTGAAGATCAAG 1281

Qy    901 TCTCTCAACTGGTATGAGGACAACAACTACAAAGCCCTCTTTGGAGGCAATGGCACTGAG 960
          || |||||||||||| ||||| |||||| ||||||||||||||||||| ||  ||| |||
Db   1282 TCCCTCAACTGGTACGAGGATAACAACTACAAAGCCCTCTTTGGAGGGAATAACACCGAG 1341

Qy    961 GAAGATGCTGAAACCTTCTATGACAACTCTACAACTCCTTACTGCAATGATTTGATGAAG 1020
          |||||  |  |  || ||||||||||| |  |||||||||| |||||||||||||||||||
Db   1342 GAAGACGTGGACACCTTCTATGACAATTCTACAACTCCTTATTGCAATGATTTGATGAAG 1401

Qy   1021 AATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGAAGCCGCTGCTCGTT 1080
          || |||||||||||||||||||| ||| ||||| ||||| || ||  || |||||| |||
Db   1402 AACTTGGAGTCTAGTCCTCTTTCTCGAATTATTTGGAAGGCACTCAAGCCACTGCTTGTT 1461
```

```
Qy    1081 GGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCATGGCTGAGGTGAAC 1140
           ||  || |  ||||||||||||||||||  ||||| ||||||||||||||||||||||||||
Db    1462 GGAAAGATTCTCTATACACCTGACACACCAGCTACAAGGCAGGTCATGGCTGAGGTGAAC 1521

Qy    1141 AAGACCTTCCAGGAACTGGCTGTGTTCCATGATCTGGAAGGCATGTGGGAGGAACTCAGC 1200
           ||||||||  |||||  |||||||||||||||||  |||||  ||||||||||  ||||||||
Db    1522 AAGACCTTTCAGGAGTTGGCTGTGTTCCATGACCTGGAGGGCATGTGGGAAGAACTCAGC 1581

Qy    1201 CCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTGTCCGGATGCTGTTG 1260
           |||  | || |||||||||||||||||||||||||||  |||||||||||||||||  |||||
Db    1582 CCCCAAATTTGGACCTTCATGGAGAACAGCCAAGAGATGGACCTTGTCCGGACGCTGTTA 1641

Qy    1261 GACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCC 1320
           |||||||| | ||||||||||| |||||||||||| ||||||||||| |||||||||| |||
Db    1642 GACAGCAGAGGCAATGACCAGTTTTGGGAACAGAAGTTGGATGGATTAGATTGGACTGCC 1701

Qy    1321 CAAGACATCGTGGCCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCT 1380
           ||||||||| |||||||| |||||||||  |||||||||||| |  |||||  ||||| |||
Db    1702 CAAGACATCATGGCCGTTTCTGGCCAAGAACCCAGAAGATGCTCAGTCCCCAAATGGCTCT 1761

Qy    1381 GTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCCGGACCATATCTCGC 1440
           ||||| |||||||||||||||||||||  ||||||||||||||||||  ||| |||||||||
Db    1762 GTGTATACCTGGAGAGAAGCTTTCAATGAGACCAACCAGGCAATCCAGACGATATCTCGA 1821

Qy    1441 TTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAGAAGTCTGGCTCATC 1500
           |||||||||||||||||||||||||||||  |||||||||  | ||||||||| |||||||
Db    1822 TTCATGGAGTGTGTCAACCTGAACAAGCTGGAACCCATTCCGACAGAAGTCAGGCTCATC 1881

Qy    1501 AACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTCACTGGA 1560
           |||||||||||||||||||||||  ||||||||||||||||||||||  || |||||||| ||
Db    1882 AACAAGTCCATGGAGCTGCTGGACGAGAGGAAGTTCTGGGCTGGCATCGTGTTCACAGGC 1941

Qy    1561 ATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGATCCGAATGGACATT 1620
           || |||||||| | |  || |||||||||||||| |||||||||||||||||| ||||||||
Db    1942 ATCACTCCAGATAGTGTGGAGCTGCCCCATCATGTAAAGTACAAGATCCGGATGGACATT 2001

Qy    1621 GACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACCCTGGTCCTCGAGCT 1680
           ||||| ||||||||| |  || |||||  ||||||||||||||||||||||||||||  |||
Db    2002 GACAACGTGGAGAGAACTAATAAGATCAAGGATGGGTACTGGGACCCTGGTCCTCGGGCT 2061

Qy    1681 GACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGGCTTCGCCTACTTGCAGGATGTGGTG 1740
           ||||| |||||| || ||||| ||||| || ||||||| ||||||||||||||||||||||
Db    2062 GACCCTTTTGAAGATATGCGCTATGTCTGGGGCGGCTTCGCCTACTTGCAGGATGTGGTG 2121

Qy    1741 GAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAACTGGTGTCTATATGCAA 1800
           || |||||| |||||||| |||||||||||| | |||||||||| |||||||||| |||||
Db    2122 GAACAGGCCATCATCAGAGTGCTGACGGGGATCTGAGAAGAAAACGGGTGTCTACGTGCAA 2181

Qy    1801 CAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGATGAGCCGGTCAATG 1860
           |||||||||||  ||||||| ||||||||||| ||||||||||| |||||||||||||||
Db    2182 CAGATGCCCTACCCCTGTTATGTTGATGACATTTTTCTGCGGGTCATGAGCCGGTCAATG 2241

Qy    1861 CCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCATCAAGGGCATCGTG 1920
           |||||||||||||| |||| ||||||| ||||| || |||||||||||||||| |||| ||
Db    2242 CCCCTCTTCATGACTCTAGCCTGGATCTACTCTGTCGCTGTGATCATCAAGAGCATTGTG 2301

Qy    1921 TATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCCTGGACAACAGCATC 1980
           |||||||||||||| | |||||||| ||||||||||||||||||||| ||||||| ||||
Db    2302 TATGAGAAGGAGGCTCGGCTGAAGGAGACCATGCGGATCATGGGTCTGGACAATGGCATC 2361

Qy    1981 CTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCCTCTTCTTGTGAGCGCTGGCCTGCTA 2040
           |||||||||||||||||| |||| |||||||||| |||| |||||||||||||||||||
Db    2362 CTCTGGTTTAGCTGGTTTGTTAGCAGCCTCATCCCTCTGCTTGTGAGCGCTGGCCTGCTG 2421

Qy    2041 GTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCAGCGTGGTGTTTGTC 2100
           |||||||||| ||||||||||||||||||||||| |||||||| |||||||||||| |||
Db    2422 GTGGTCATCTTGAAGTTAGGAAACCTGCTGCCCTATAGTGACCCCAGCGTGGTGTTCGTC 2481

Qy    2101 TTCCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGATTAGCACACTCTTC 2160
           ||||||| ||||||||| |||||||||||||| ||||||||||||| ||||||||| ||||
Db    2482 TTCCTGTCTGTGTTTGCCATGGTGACCATCCTACAGTGCTTCCTCATTAGCACGCTCTTC 2541
```

```
Qy    2161 TCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCACGCTGTACCTGCCC 2220
           ||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2542 TCCCGTGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCACGCTGTACCTGCCC 2601

Qy    2221 TACGTCCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCAAGATCTTCGCTAGC 2280
           || || |||||| || || ||||||||||| ||||||||| | ||||||||||| ||||||
Db    2602 TATGTGCTGTGCGTAGCCTGGCAGGACTATGTGGGCTTCTCCATCAAGATCTTTGCTAGC 2661

Qy    2281 CTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCCTTTTTGAGGAGCAG 2340
           ||||||||||||||||||||||||| || |||||||||||| ||||||||  ||||||||
Db    2662 CTGCTGTCTCCTGTGGCTTTTGGATTCGGCTGTGAGTATTTCGCCCTTTTCGAGGAGCAA 2721

Qy    2341 GGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGGAAGATGGCTTCAAT 2400
           || || || || || || ||||||| || |||||||| || ||||||||| || |||||||
Db    2722 GGTATCGGGGTCCAATGGGACAATCTCTTTGAGAGCCCGGTGGAGGAGGACGGCTTCAAT 2781

Qy    2401 CTCACCACTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATGGGGTGATGACCTGG 2460
           |||||||| | || |||||||||||||| ||||||||| |||||||| |||||| ||||
Db    2782 CTCACCACTGCAGTGTCCATGATGCTCTTTGACACCTTTCTCTATGGCGTGATGACATGG 2841

Qy    2461 TACATTGAGGCTGTCTTTCCAGGCCAGTACGGAATTCCCAGGCCCTGGTATTTTCCTTGC 2520
           ||||| || || |||||||||||| ||||| ||||| ||||||||||||||||||||||||
Db    2842 TACATCGAAGCCGTCTTTCCAGGACAGTATGGAATTCCCAGGCCCTGGTATTTTCCTTGT 2901

Qy    2521 ACCAAGTCCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACCCTGGTTCCAACCAG 2580
           |||||||| |||||||||||| ||||||| |||||||||||||||||||||||| ||||
Db    2902 ACCAAGTCATACTGGTTTGGTGAGGAAATTGATGAGAAGAGCCACCCTGGTTCCAGCCAG 2961

Qy    2581 AAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGAAGCTGGGCGTGTCC 2640
           ||| || | ||||||||||||||||||||| ||||||| ||  ||| |||||| ||||||
Db    2962 AAGGGAGTGTCAGAAATCTGCATGGAAGAGGAACCCACTCATCTGAGGCTGGGGGTGTCC 3021

Qy    2641 ATTCAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTCGATGGCCTGGCA 2700
           ||||||||||||||| || || !||||||||||| |||||||| ||||| ||||| ||||
Db    3022 ATTCAGAACCTGGTGAAGGTTTACCGAGATGGCATGAAGGTTGCTGTGGATGGCTTGGCG 3081

Qy    2701 CTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATGGAGCGGGGAAGACG 2760
           || || ||||| || ||||||||| |||||||||||||||||||||||||||| ||||||
Db    3082 CTCAACTTTTACGAAGGCCAGATTACCTCCTTCCTGGGCCACAATGGAGCAGGGAAGACC 3141

Qy    2761 ACCACCATGTCAATCCTGACCGGGTTGTTCCCCCCGACCTCGGGCACCGCCTACATCCTG 2820
           |||||||||||||||| |||||| ||| |||| !|||| || || ||||| |||||||||||
Db    3142 ACCACCATGTCAATACTGACTGGGCTGTTTCCCCCAACTTCTGGCACGGCCTACATCCTG 3201

Qy    2821 GGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTCTGTCCCCAG 2880
           || || |||||||||| ||||||||||| |||||||||||||||||||| |||||||||||
Db    3202 GGGAAGGACATTCGCTCGGAGATGAGCTCCATCCGGCAGAACCTGGGAGTCTGTCCCCAG 3261

Qy    2881 CATAACGTGCTGTTTGACATGCTGACTGTCGAAGAACACATCTGGTTCTATGCCCGCTTG 2940
           ||||| |||||||||||||||||||||||||||| |||||||||||||||||||  ||| |
Db    3262 CATAATGTGCTGTTTGACATGCTGACTGTCGAAGAGCACATCTGGTTCTATGCGCGCCTA 3321

Qy    2941 AAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGGCCCTGGATGTTGGT 3000
           || |||||||| |||||||||||||| || ||||||||||||||||||||||||||||||
Db    3322 AAGGGGCTCTCAGAGAAGCACGTGAAAGCAGAGATGGAGCAGATGGCCCTGGATGTTGGC 3381

Qy    3001 TTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCAGGTGGAATGCAGAGAAAG 3060
           || ||  | |||||||||||||||||| || ||||| |||||| |||||||||||||||||
Db    3382 TTACCCCCGAGCAAGCTGAAAAGCAAAACGAGTCAGCTCTCAGGTGGGATGCAGAGAAAG 3441

Qy    3061 CTATCTGTGGCCTTGGCCTTTGTCGGGGGATCTAAGGTTGTCATTCTGGATGAACCCACA 3120
           || |||||||||||||||||| || || |||||||||||||||||||||| || |||||||
Db    3442 CTGTCTGTGGCCTTGGCCTTCGTGGGTGGATCCAAGGTTGTCATTCTGGACGAGCCCACA 3501

Qy    3121 GCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTGCTGCTGAAATACCGACAA 3180
           || || |||||||||| ||||| |||||||||||||||| || ||||| |||||||| |||
Db    3502 GCCGGGGTGGACCCGTACTCTCGCAGGGGAATATGGGAACTCCTGCTAAAATACCGGCAA 3561

Qy    3181 GGCCGCACCATTATTCTCTCTACACACCACATGGATGAAGCGGACGTCCTGGGGGACAGG 3240
           ||||||||||||||| | |||||||||||||||||| ||||  ||| ||| |||||||||
```

```
Db    3562 GGCCGCACCATTATTTTGTCTACACACCACATGGACGAAGCTGACATCCTTGGGGACAGA 3621

Qy    3241 ATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTCTGAAGAAC 3300
           ||||||||||| ||||||||||||||||| ||||||||||||||||||||||| |||| |||
Db    3622 ATTGCCATCATTTCCCATGGGAAGCTGTGTTGTGTGGGCTCCTCCCTGTTTTTTGAAAAAC 3681

Qy    3301 CAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGT 3360
           ||| ||||||| || ||||| |||||| |||| |||||||||||||||||||| ||||||||
Db    3682 CAGTTGGGAACGGGTTACTATCTGACCCTGGTTAAGAAAGATGTGGAATCGTCCCTCAGT 3741

Qy    3361 TCCTGCAGAAACAGTAGTAGCACTGTGTCATACCTGAAAAAGGAGGACAGTGTTTCTCAG 3420
           ||||||||||||||||| ||||| |||||| | ||||||||||||||||||||||||||||
Db    3742 TCCTGCAGAAACAGTAGCAGCACCGTGTCTTGTCTGAAAAAGGAGGACAGTGTTTCTCAG 3801

Qy    3421 AGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGATGTC 3480
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    3802 AGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAAAGTGACACGCTGACCATCGATGTC 3861

Qy    3481 TCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCGGCTGGTGGAAGACATA 3540
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db    3862 TCTGCTATCTCCAACCTCATCAGGAAGCACGTGTCTGAAGCCCGGCTGGTGGAGGACATT 3921

Qy    3541 GGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTTGTGGAA 3600
           ||||| ||||||||||||||||||||||| || ||||| || |||||||||||||||||||
Db    3922 GGGCACGAGCTGACCTATGTGCTGCCGTACGAAGCCGCGAAGGAGGGAGCCTTTGTGGAA 3981

Qy    3601 CTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCA 3660
           ||||| |||||||||||||||||||||||||||||||||||| || |||||||||||||||
Db    3982 CTCTTCCATGAGATTGATGACCGGCTCTCAGACCTGGGCATCTCCAGTTATGGCATCTCG 4041

Qy    3661 GAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAG 3720
           ||||| |||||||||||||||||||||||| |||||| ||||||||| ||||||||||||||
Db    4042 GAGACCACCCTGGAAGAAATATTCCTCAAAGTGGCTGAAGAGAGCGGGGTGGATGCTGAG 4101

Qy    3721 ACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCGGGGACAAGCAGAGC 3780
           |||||||||||||| ||||||||||||||||||||| |||||||||||||||||||||||||
Db    4102 ACCTCAGATGGTACTTTGCCAGCAAGACGAAACAGACGGGCCTTCGGGGACAAGCAGAGC 4161

Qy    3781 TGTCTTCGCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATTCTGACATAGACCCA 3840
           ||||| | ||| || || |||||||||||||| |||||| ||||| ||||||||||||||
Db    4162 TGTCTGCACCCATTTACGGAAGATGATGCTGTTGATCCCAATGACTCTGACATAGACCCA 4221

Qy    3841 GAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGTCCTACCAGGTGAAA 3900
           ||||||| ||||| ||| |||||||||||||||||| ||||||||| |||||||| ||||
Db    4222 GAATCCAGGGAGACCGACCTGCTCAGTGGGATGGACGGCAAAGGCTCCTACCAGCTGAAG 4281

Qy    3901 GGCTGGAAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGACTGCTAATTGCC 3960
           ||||||||||| || ||||||||||||||||||| |||||||||||||| ||||| ||||||
Db    4282 GGCTGGAAACTCACCCAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGGCTGCTGATTGCC 4341

Qy    3961 AGACGGAGTCGGAAAGGATTTTTTGCTCAGATTGTCTTGCCAGCTGTGTTTGTCTGCATT 4020
           ||||||||| |||||| || || |||||||||||||||| ||||||||| ||||| ||||||
Db    4342 AGACGGAGCCGGAAGGGTTTCTTTGCTCAGATTGTCCTGCCAGCTGTCTTTGTTTGCATT 4401

Qy    4021 GCCCTTGTGTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCCAGCCTGGAACTTCAG 4080
           ||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4402 GCCCTGGTCTTCAGCCTGATTGTGCCACCCTTTGGCAAGTACCCCAGCCTGGAACTTCAG 4461

Qy    4081 CCCTGGATGTACAACGAACAGTACACATTTGTCAGCAATGATGCTCCTGAGGACACGGGA 4140
           |||||||||||| || || |||||| ||||||||||||||| |||||||| ||||||| ||
Db    4462 CCCTGGATGTATAATGAGCAGTATACATTTGTCAGTAATGATGCTCCCGAGGACATGGGC 4521

Qy    4141 ACCCTGGAACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAA 4200
           |||| |||||| | || || || |||||||| ||||| |||||||||||||||||||||||
Db    4522 ACCCAGGAACTCCTGAATGCTCTGACCAAAGATCCAGGCTTTGGGACCCGCTGTATGGAA 4581

Qy    4201 GGAAACCCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGTGGACCACTGCCCCA 4260
           |||||||||||||||||| || || ||| ||| ||||||||| || |||||||| |||
Db    4582 GGAAACCCAATCCCAGATACCCCTTGCTTGGCTGGGGAGGAGGACTGGACCATCAGCCCC 4641

Qy    4261 GTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACTGGACAATGCAGAACCCTTCA 4320
```

```
         || ||||||||  ||||  |||||||||||||||||||| |||||||| ||| |||||||| |||
Db   4642 GTCCCCCAGAGCATCGTGGACCTCTTCCAGAATGGAAACTGGACCATGAAGAACCCCTCA 4701

Qy   4321 CCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTGTGTGTCCCCCAGGG 4380
          |||||  ||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   4702 CCTGCGTGCCAGTGTAGCAGTGACAAAATCAAGAAGATGCTGCCTGTGTGTCCCCCAGGG 4761

Qy   4381 GCAGGGGGGCTGCCTCCTCCAC 4402
          |||||||||||||| |||||| |
Db   4762 GCAGGGGGGCTGCCACCTCCTC 4783
```

SEQ ID NO: 2

Result No.	Score	Query Match	Length	DB	ID	Description
1	11469	100.0	2201	21	AAY79380	Human ATP binding
2	11469	100.0	2201	22	AAE13021	Human ATP binding
3	11469	100.0	2201	22	AAM50227	Human ATP binding
4	11469	100.0	2201	23	ABP65164	Hypoxia-regulated
5	11469	100.0	2261	22	AAE13022	Human ATP binding
6	11469	100.0	2261	22	AAM50228	Human ATP binding
7	11469	100.0	2261	22	AAU02176	Human ABC1. Homo
8	11469	100.0	2261	22	AAU02177	Human ABC1 mutant
9	11469	100.0	2261	23	ABB83111	Human ABCA1 transp
10	11469	100.0	2261	23	ABP52092	Homo sapiens ABC t
11	11469	100.0	2261	23	AAE23000	Human ABC1 full-le
12	11468	100.0	2261	23	ABB83117	Polymorphic human
13	11468	100.0	2261	23	ABB83124	Polymorphic human
14	11466	100.0	2261	22	AAU02183	Human ABC1 mutant
15	11466	100.0	2261	22	AAU02188	Human ABC1 mutant
16	11466	100.0	2261	23	ABB83115	Polymorphic human
17	11466	100.0	2261	23	ABB83116	Polymorphic human
18	11466	100.0	2261	23	ABB83119	Polymorphic human
19	11466	100.0	2261	23	ABB83122	Polymorphic human
20	11466	100.0	2261	23	ABB83123	Polymorphic human
21	11465	100.0	2261	22	AAU02189	Human ABC1 mutant
22	11464	100.0	2261	23	ABB83121	Polymorphic human
23	11463	99.9	2261	22	AAU02181	Human ABC1 mutant
24	11462	99.9	2261	22	AAM78550	Human protein SEQ
25	11462	99.9	2263	22	ABB11956	Human ABCA1 homolo
26	11462	99.9	2263	22	AAM79534	Human protein SEQ
27	11461	99.9	2261	22	AAU02182	Human ABC1 mutant
28	11461	99.9	2261	22	AAU02186	Human ABC1 mutant
29	11459	99.9	2261	23	ABB83118	Polymorphic human
30	11458	99.9	2261	23	ABB83120	Polymorphic human
31	11440	99.7	2261	21	AAB38082	Human ABC1 cholest
32	11440	99.7	2261	22	AAB71749	Human ABC1 protein
33	11440	99.7	2261	24	ABU11899	Human ATP-binding
34	11439	99.7	2261	23	ABB81578	Human ABC-A-1-1 pr
35	11437	99.7	2261	21	AAB38109	Human ABC1 cholest
36	11437	99.7	2261	21	AAB38111	Human ABC1 cholest
37	11437	99.7	2261	21	AAB38114	Human ABC1 cholest
38	11437	99.7	2261	21	AAB38115	Human ABC1 cholest
39	11437	99.7	2261	21	AAB38117	Human ABC1 cholest

No.	Score	Match	Length	DB	ID	Description
1	6909	60.2	1375	3	US-08-665-259-26	Sequence 26, Appl
2	6909	60.2	1375	3	US-08-762-500-26	Sequence 26, Appl
3	3129.5	27.3	1457	3	US-08-665-259-27	Sequence 27, Appl
4	3129.5	27.3	1457	3	US-08-762-500-27	Sequence 27, Appl
5	2635.5	23.0	1684	3	US-08-665-259-25	Sequence 25, Appl
6	2635.5	23.0	1684	3	US-08-762-500-25	Sequence 25, Appl
7	2635.5	23.0	1704	3	US-08-762-500-75	Sequence 75, Appl

```
 8     359    3.1    315   4   US-09-328-352-4388    Sequence 4388, Ap
 9     354    3.1    332   4   US-09-107-532A-3752   Sequence 3752, Ap
10   352.5    3.1    335   4   US-09-252-991A-20837  Sequence 20837, A
11     352    3.1    322   4   US-09-107-532A-4662   Sequence 4662, Ap
12     346    3.0    589   4   US-09-328-352-7592    Sequence 7592, Ap
13   336.5    2.9   1280   2   US-08-583-276-19      Sequence 19, Appl
14   335.5    2.9    291   4   US-09-107-532A-4205   Sequence 4205, Ap
15   335.5    2.9   1280   4   US-09-767-594-2       Sequence 2, Appli
16   335.5    2.9   1280   6   5206352-4             Patent No. 5206352
17     333    2.9   1279   2   US-08-784-649A-2      Sequence 2, Appli
18     332    2.9    402   4   US-09-107-532A-5360   Sequence 5360, Ap
19   330.5    2.9   1280   2   US-08-752-447-2       Sequence 2, Appli
```

RESULT 1
US-08-665-259-26
; Sequence 26, Application US/08665259
; Patent No. 6028173
;   GENERAL INFORMATION:
;   APPLICANT: Landes, Gregory M.
;   APPLICANT: Burn, Timothy C.
;   APPLICANT: Connors, Timothy D.
;   APPLICANT: Dackowski, William R.
;   APPLICANT: Van Raay, Terence J.
;   APPLICANT: Klinger, Katherine W.
;   TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
;   TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
;   NUMBER OF SEQUENCES: 73
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: GENZYME CORPORATION
;   STREET: One Mountain Road
;   CITY: Framingham
;   STATE: Massachusetts
;   COUNTRY: United States of America
;   ZIP: 01701
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/665,259
;   FILING DATE: 17-JUN-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Dugan, Deborah A.
;   REGISTRATION NUMBER: 37,315
;   REFERENCE/DOCKET NUMBER: IG5-9.1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (508) 872-8400
;   TELEFAX: (508) 872-5415
;   INFORMATION FOR SEQ ID NO: 26:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1375 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
US-08-665-259-26

  Query Match   60.2%; Score 6909; DB 3; Length 1375;
  Best Local Similarity 96.9%; Pred. No. 0;
  Matches 1332; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

Qy   827 CMEEEPTHLKLGVSIQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSIL 886
            |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CMEEEPTHLRLGVSIQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSIL 60

Qy   887 TGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEK 946
            |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
```

```
Db       61 TGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEK 120

Qy      947 HVKAEMEQMALDVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPY 1006
            ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      121 HVKAEMEQMALDVGLPPSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPY 180

Qy     1007 SRRGIWELLLKYRQGRTIILSTHHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGY 1066
            |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db      181 SRRGIWELLLKYRQGRTIILSTHHMDEADILGDRIAIISHGKLCCVGSSLFLKNQLGTGY 240

Qy     1067 YLTLVKKDVESSLSSCRNSSSTVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNL 1126
            ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      241 YLTLVKKDVESSLSSCRNSSSTVSCLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNL 300

Qy     1127 IRKHVSEARLVEDIGHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEE 1186
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 IRKHVSEARLVEDIGHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEE 360

Qy     1187 IFLKVAEESGVDAETSDGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETD 1246
            |||||||||||||||||||||||||||||||||||| ||||||  |||||||||||||||
Db      361 IFLKVAEESGVDAETSDGTLPARRNRRAFGDKQSCLHPFTEDDAVDPNDSDIDPESRETD 420

Qy     1247 LLSGMDGKGSYQVKGWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSL 1306
            ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 LLSGMDGKGSYQLKGWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSL 480

Qy     1307 IVPPFGKYPSLELQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPD 1366
            ||||||||||||||||||||||||||||||||| || |||||||||||||||||||||||
Db      481 IVPPFGKYPSLELQPWMYNEQYTFVSNDAPEDMGTQELLNALTKDPGFGTRCMEGNPIPD 540

Qy     1367 TPCQAGEEEWTTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPP 1426
            ||| || ||:|| :|||| :||||||||||| :||||||||||||||||||||||||||||
Db      541 TPCLAGEEDWTISPVPQSIVDLFQNGNWTMKNPSPACQCSSDKIKKMLPVCPPGAGGLPP 600

Qy     1427 PQRKQNTADILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALP 1486
            ||||| |||||| :||||||||||||||||||||||||||||||||||||||||:|||
Db      601 PQRKQKTADILQNLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNSQALP 660

Qy     1487 PSQEVNDATKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFL 1546
            || |||||:|||||| || ||: |||||:|||||| ||| :|||||||||||||||||||
Db      661 PSHEVNDAIKQMKKLLKLTKDTSADRFLSSLGRFMAGLDTKNNVKVWFNNKGWHAISSFL 720

Qy     1547 NVINNAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMS 1606
            ||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db      721 NVINNAILRANLQKGENPSQYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMS 780

Qy     1607 FVPASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ 1666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 FVPASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ 840

Qy     1667 KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATF 1726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATF 900

Qy     1727 VLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW 1786
            ||||||:||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 VLELFTNNKLNDINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW 960

Qy     1787 DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGG 1846
            ||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||
Db      961 DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVKAKLPPLNDEDEDVRRERQRILDGG 1020

Qy     1847 GQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLTGDTTVT 1906
            ||||||||||||||||||||||||||||:|||||||||||||||||||:||||||||| ||
Db     1021 GQNDILEIKELTKIYRRKRKPAVDRICIGIPPGECFGLLGVNGAGKSTTFKMLTGDTPVT 1080

Qy     1907 RGDAFLNRNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGE 1966
            |||||||:||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     1081 RGDAFLNKNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKFGE 1140

Qy     1967 WAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNC 2026
```

```
       ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     1141 WAIRKLGLVKYGEKYASNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNC 1200

Qy     2027 ALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIA 2086
            |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 ALSIVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIA 1260

Qy     2087 GSNPDLKPVQDFFGLAFPGSVPKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDY 2146
            ||||||||||:|||||||||| |||||||||||||||||||||||||||||||||||||
Db     1261 GSNPDLKPVQEFFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDY 1320

Qy     2147 SVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 2201
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 SVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 1375
```

RESULT 2
US-08-762-500-26
; Sequence 26, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IG5-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1375 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-762-500-26

    Query Match   60.2%; Score 6909; DB 3; Length 1375;
```

Best Local Similarity   96.9%; Pred. No. 0;
Matches 1332; Conservative   21; Mismatches   22; Indels   0; Gaps   0;

```
Qy       827 CMEEEPTHLKLGVSIQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSIL 886
             |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CMEEEPTHLRLGVSIQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSIL 60

Qy       887 TGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEK 946
             |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db        61 TGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEK 120

Qy       947 HVKAEMEQMALDVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPY 1006
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       121 HVKAEMEQMALDVGLPPSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPY 180

Qy      1007 SRRGIWELLLKYRQGRTIILSTHHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGY 1066
             |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db       181 SRRGIWELLLKYRQGRTIILSTHHMDEADILGDRIAIISHGKLCCVGSSLFLKNQLGTGY 240

Qy      1067 YLTLVKKDVESSLSSCRNSSSTVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNL 1126
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       241 YLTLVKKDVESSLSSCRNSSSTVSCLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNL 300

Qy      1127 IRKHVSEARLVEDIGHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEE 1186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 IRKHVSEARLVEDIGHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEE 360

Qy      1187 IFLKVAEESGVDAETSDGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETD 1246
             |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db       361 IFLKVAEESGVDAETSDGTLPARRNRRAFGDKQSCLHPFTEDDAVDPNDSDIDPESRETD 420

Qy      1247 LLSGMDGKGSYQVKGWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSL 1306
             |||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 LLSGMDGKGSYQLKGWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSL 480

Qy      1307 IVPPFGKYPSLELQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPD 1366
             |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||
Db       481 IVPPFGKYPSLELQPWMYNEQYTFVSNDAPEDMGTQELLNALTKDPGFGTRCMEGNPIPD 540

Qy      1367 TPCQAGEEEWTTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPP 1426
             ||| ||||:|| :||||:|:|||||||||||:||||||||||||||||||||||||||||
Db       541 TPCLAGEEDWTISPVPQSIVDLFQNGNWTMKNPSPACQCSSDKIKKMLPVCPPGAGGLPP 600

Qy      1427 PQRKQNTADILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALP 1486
             ||||| ||||||:|||||||||||||||||||||||||||||||||||||||||:|||||
Db       601 PQRKQKTADILQNLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNSQALP 660

Qy      1487 PSQEVNDATKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFL 1546
             || ||||| ||||| ||| ||:|||||:||||| |||:||||:|||||||||||||||||
Db       661 PSHEVNDAIKQMKKLLKLTKDTSADRFLSSLGRFMAGLDTKNNVKVWFNNKGWHAISSFL 720

Qy      1547 NVINNAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMS 1606
             |||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||||
Db       721 NVINNAILRANLQKGENPSQYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMS 780

Qy      1607 FVPASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ 1666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 FVPASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQ 840

Qy      1667 KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATF 1726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 KSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATF 900

Qy      1727 VLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW 1786
             ||||||:||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 VLELFTNNKLNDINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW 960

Qy      1787 DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGG 1846
             ||||||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||
Db       961 DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVKAKLPPLNDEDEDVRRERQRILDGG 1020
```

```
Qy   1847 GQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLTGDTTVT 1906
          |||||||||||||||||||||||||||||:|||||||||||||||||||||||:|||||||| ||
Db   1021 GQNDILEIKELTKIYRRKRKPAVDRICIGIPPGECFGLLGVNGAGKSTTFKMLTGDTPVT 1080

Qy   1907 RGDAFLNRNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGE 1966
          |||||||:||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1081 RGDAFLNKNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKFGE 1140

Qy   1967 WAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNC 2026
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1141 WAIRKLGLVKYGEKYASNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNC 1200

Qy   2027 ALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIA 2086
          |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ALSIVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIA 1260

Qy   2087 GSNPDLKPVQDFFGLAFPGSVPKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDY 2146
          ||||||||||:|||||||||||||:||||||||||||||||||||||||||||||||||||
Db   1261 GSNPDLKPVQEFFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDY 1320

Qy   2147 SVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 2201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 SVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 1375
```

                              SUMMARIES

		% Query				
Result No.	Score	Match	Length	DB	ID	Description
1	10906	95.1	2201	2	A54774	ATP binding casset
2	3338.5	29.1	1529	2	A59189	ATP-binding casset
3	3129.5	27.3	1472	2	B54774	ATP binding casset
4	2638.5	23.0	1704	2	S71363	probable ATP-bindi
5	2635.5	23.0	1704	2	A59188	ATP-binding casset
6	2061	18.0	1802	2	T33783	hypothetical prote
7	1975	17.2	1816	2	A84845	probable ABC trans
8	1854.5	16.2	1447	2	T15200	hypothetical prote
9	1792	15.6	1317	2	C88925	protein F33E11.4 [
10	1535.5	13.4	1758	2	F88559	protein C48B4.4b [
11	1528.5	13.3	1704	2	T42749	ATP-binding casset
12	1526	13.3	1767	2	S60124	transport protein
13	1393	12.1	1246	2	T00826	hypothetical prote
14	1150.5	10.0	1564	2	T27121	hypothetical prote
15	1006.5	8.8	373	2	T47150	hypothetical prote
16	1005.5	8.8	1431	2	T22748	hypothetical prote
17	846	7.4	269	2	T46467	hypothetical prote
18	786.5	6.9	1011	2	T07712	probable ABC-type
19	695	6.1	900	2	T07717	probable ABC-type

```
RESULT 1
A54774
ATP binding cassette transporter ABC1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 02-Feb-2001
C;Accession: A54774
R;Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A;Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A;Reference number: A54774; MUID:94375008; PMID:8088782
A;Accession: A54774
A;Molecule type: mRNA
A;Residues: 1-2201 <LUC>
A;Cross-references: GB:X75926; NID:g495256; PIDN:CAA53530.1; PID:g495257
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP; duplication; nucleotide binding; P-loop
F;856-1047/Domain: ATP-binding cassette homology <ABC1>
F;873-880/Region: nucleotide-binding motif A (P-loop)
```

F;1869-2060/Domain: ATP-binding cassette homology <ABC2>
F;1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match   95.1%; Score 10906; DB 2; Length 2201;
Best Local Similarity   94.8%; Pred. No. 0;
Matches 2087; Conservative   54; Mismatches   60; Indels   0; Gaps   0;

```
Qy       1 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDARRLLLYSQKDT 60
           |||||||||||||||||||||||||||||||||||||||||||| :|||||| :|||||| :||
Db       1 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVSRLFSDAQRLLLYSQRDT 60

Qy      61 SMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILH 120
           |:||| |||| |:||| :||||||||||||||||| ||||| :|||| :|: :| | 
Db      61 SIKDMHKVLRMLRQIKHPNSNLKLQDFLVDNETFSGFLQHNLSLPRSTVDSLLQXNVGLQ 120

Qy     121 KVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLPREKLAAAERVLRSNMDILKPIL 180
           |||||||||||| |||||||| ||:||||| ||| |||||||:|| ||||||| |||||||:::
Db     121 KVFLQGYQLHLASLCNGSKLEEIIQLGDAEVSALCGLPRKKLDAAERVLRYNMDILKPVV 180

Qy     181 RTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMRSWSDMRQEVMFLTNVNSSSSSTQI 240
           ||||| |:: ||||| || ||| ||||||| :|||||||||||||||||||||||||||||
Db     181 TKLNSTSHLPTQHLAEATTVLLDSLGGLAQELFSTKSWSDMRQEVMFLTNVNSSSSSTQI 240

Qy     241 YQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNGTEEDAETFYDNSTTPYCNDLMK 300
           |||||||||||||||||||||||||||||||||||||||| |||| :||||||||||||||||
Db     241 YQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNNTEEDVDTFYDNSTTPYCNDLMK 300

Qy     301 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTFQELAVFHDLEGMWEELS 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTFQELAVFHDLEGMWEELS 360

Qy     361 PKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTAQDIVAFLAKHPEDVQSSNGS 420
           |:||||||||||||||||:|||| ||||| |||| :||||||||||:|||||:|||||| |||
Db     361 PQIWTFMENSQEMDLVRTLLDSRGNDQFWEQKLDGLDWTAQDIMAFLAKNPEDVQSPNGS 420

Qy     421 VYTWREAFNETNQAIRTISRFMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTG 480
           ||||||||||||||||:||||||||||||||||| |||| ||||||||||||||||||||||
Db     421 VYTWREAFNETNQAIQTISRFMECVNLNKLEPIPTEVRLINKSMELLDERKFWAGIVFTG 480

Qy     481 ITPGSIELPHHVKYKIRMDIDNVERTNKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVV 540
           ||| |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ITPDSVELPHHVKYKIRMDIDNVERTNKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVV 540

Qy     541 EQAIIRVLTGTEKKTGVYMQQMPYPCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIV 600
           |||||||||| :|||||||:|||||||||||||||||||||||||||||||||||||||| ||
Db     541 EQAIIRVLTGSEKKTGVYVQQMPYPCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKSIV 600

Qy     601 YEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV 660
           ||||||||||||||||||| |||||||:|||||||||||||||||||||||||||||||||
Db     601 YEKEARLKETMRIMGLDNGILWFSWFVSSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV 660

Qy     661 FLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFAS 720
           ||||||:|||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db     661 FLSVFAMVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFSIKIFAS 720

Qy     721 LLSPVAFGFGCEYFALFEEQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTW 780
           |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db     721 LLSPVAFGFGCEYFALFEEQGIGVQWDNLFESPVEEDGFNLTTAVSMMLFDTFLYGVMTW 780

Qy     781 YIEAVFPGQYGIPRPWYFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVS 840
           ||||||||||||||||||||||||||||| ||||||||:|| :||||||||||||:||||||
Db     781 YIEAVFPGQYGIPRPWYFPCTKSYWFGEEIDEKSHPGSSQKGVSEICMEEEPTHLRLGVS 840

Qy     841 IQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNAGKTTTMSILTGLFPPTSGTAYIL 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 IQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNAGKTTTMSILTGLFPPTSGTAYIL 900

Qy     901 GKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVG 960
           |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GKDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVG 960
```

```
Qy    961 LPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQ 1020
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 LPPSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQ 1020

Qy   1021 GRTIILSTHHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLS 1080
          |||||||||||||||| :||||||||||||||||||||||||||||||||||||||||||
Db   1021 GRTIILSTHHMDEADILGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLS 1080

Qy   1081 SCRNSSSTVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDI 1140
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 SCRNSSSTVSCLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDI 1140

Qy   1141 GHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAE 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAE 1200

Qy   1201 TSDGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK 1260
          ||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||:|
Db   1201 TSDGTLPARRNRRAFGDKQSCLHPFTEDDAVDPNDSDIDPESRETDLLSGMDGKGSYQLK 1260

Qy   1261 GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQ 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQ 1320

Qy   1321 PWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEWTTAP 1380
          |||||||||||||||||| || |||||||||||||||||||||||||| ||||:|| :|
Db   1321 PWMYNEQYTFVSNDAPEDMGTQELLNALTKDPGFGTRCMEGNPIPDTPCLAGEEDWTISP 1380

Qy   1381 VPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTADILQDL 1440
          |||:|:|:||||||||:|||||||||||||||||||||||||||||||||||| |||||:|
Db   1381 VPQSIVDLFQNGNWTMKNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQKTADILQNL 1440

Qy   1441 TGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVNDATKQMKK 1500
          |||||||||||||||||||||||||||||||||||||||:||||||| ||||| |||||
Db   1441 TGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNSQALPPSHEVNDAIKQMKK 1500

Qy   1501 HLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQK 1560
          ||| ||:||||||:|||||||:||||:||||||||||||||||||||||||||||||||||
Db   1501 LLKLTKDTSADRFLSSLGRFMAGLDTKNNVKVWFNNKGWHAISSFLNVINNAILRANLQK 1560

Qy   1561 GENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMSFVPASFVVFLIQER 1620
          |||||  ||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1561 GENPSQYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFVPASFVVFLIQER 1620

Qy   1621 VSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLAL 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 VSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLAL 1680

Qy   1681 LLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTDNKLNNIN 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||:||||:|||
Db   1681 LLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTNNKLNDIN 1740

Qy   1741 DILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSWDLVGRNLFAMAVEG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 DILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSWDLVGRNLFAMAVEG 1800

Qy   1801 VVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGGGQNDILEIKELTKI 1860
          |||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||
Db   1801 VVFFLITVLIQYRFFIRPRPVKAKLPPLNDEDEDVRRERQRILDGGGQNDILEIKELTKI 1860

Qy   1861 YRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLTGDTTVTRGDAFLNRNSILSN 1920
          |||||||||||||:|||||||||||||||||||:|||||||||| |||||||||:||||||
Db   1861 YRRKRKPAVDRICIGIPPGECFGLLGVNGAGKSTTFKMLTGDTPVTRGDAFLNKNSILSN 1920

Qy   1921 IHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGEWAIRKLGLVKYGEK 1980
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1921 IHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKFGEWAIRKLGLVKYGEK 1980

Qy   1981 YAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVVLT 2040
          ||  |||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db   1981 YASNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNCALSIVKEGRSVVLT 2040
```

```
Qy    2041 SHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQDFFG 2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db    2041 SHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQEFFG 2100

Qy    2101 LAFPGSVPKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDYSVSQTTLDQVFVNF 2160
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 LAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDYSVSQTTLDQVFVNF 2160

Qy    2161 AKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 2201
           ||||||||||||||||||||||||||||||||||||||||
Db    2161 AKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV 2201
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	11466	100.0	2261	1	ABC1_HUMAN	O95477 homo sapien
2	10906	95.1	2261	1	ABC1_MOUSE	P41233 mus musculu
3	5689.5	49.6	2273	1	ABCR_HUMAN	P78363 homo sapien
4	4131	36.0	2436	1	ABC2_HUMAN	Q9bzc7 homo sapien
5	3989.5	34.8	2434	1	ABC2_MOUSE	P41234 mus musculu
6	2635.5	23.0	1704	1	ABC3_HUMAN	Q99758 homo sapien
7	1528.5	13.3	1704	1	CED7_CAEEL	P34358 caenorhabdi
8	411	3.6	330	1	DRRA_STRPE	P32010 streptomyce
9	380.5	3.3	343	1	NODI_RHISN	P55476 rhizobium s
10	366	3.2	304	1	NODI_RHIS3	P72335 rhizobium s
11	347	3.0	308	1	YADG_ECOLI	P36879 escherichia
12	347	3.0	335	1	NDI2_RHIME	Q8gnh6 rhizobium m
13	344.5	3.0	340	1	NODI_RHILO	P23703 rhizobium l
14	343.5	3.0	347	1	NODI_RHIGA	P50332 rhizobium g
15	335.5	2.9	1280	1	MDR1_HUMAN	P08183 homo sapien
16	331.5	2.9	355	1	NDI1_RHIME	O52618 rhizobium m
17	329.5	2.9	578	1	YBHF_ECOLI	P75776 escherichia
18	327	2.9	894	1	YHIH_ECOLI	P37624 escherichia
19	327	2.9	1276	1	MDR3_MOUSE	P21447 mus musculu

```
RESULT 2
ABC1_MOUSE
ID   ABC1_MOUSE      STANDARD;      PRT;   2261 AA.
AC   P41233;
DT   01-FEB-1995 (Rel. 31, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette
DE   transporter 1) (ATP-binding cassette 1) (ABC-1).
GN   ABCA1 OR ABC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=Macrophage;
RX   MEDLINE=94375008; PubMed=8088782;
RA   Luciani M.F., Denizot F., Savary S., Mattei M.-G., Chimini G.;
RT   "Cloning of two novel ABC transporters mapping on human chromosome
RT   9.";
RL   Genomics 21:150-159(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=21251004; PubMed=11352567;
RA   Qiu Y., Cavelier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT   "Human and mouse ABCA1 comparative sequencing and transgenesis
```

RT   studies revealing novel regulatory sequences.";
RL   Genomics 73:66-76(2001).
CC   -!- FUNCTION: CAMP-DEPENDENT AND SULFONYLUREA-SENSITIVE ANION
CC   TRANSPORTER. KEY GATEKEEPER INFLUENCING INTRACELLULAR CHOLESTEROL
CC   TRANSPORT (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN ADULT TISSUES. HIGHEST
CC   LEVELS ARE FOUND IN PREGNANT UTERUS AND UTERUS.
CC   -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC   EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC   ATP BINDING CASSETTE (ABC) DOMAIN.
CC   -!- PTM: Phosphorylation on Ser-2054 regulates phospholipid efflux (By
CC   similarity).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. ABCA SUBFAMILY.
CC   ---

DR   EMBL; X75926; CAA53530.1; ALT_INIT.
DR   EMBL; AF287263; AAG39073.1; ALT_INIT.
DR   MGD; MGI:99607; Abca1.
DR   GO; GO:0008203; P:cholesterol metabolism; IDA.
DR   GO; GO:0030301; P:cholesterol transport; IDA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 2.
DR   ProDom; PD000006; ABC_transporter; 2.
DR   SMART; SM00382; AAA; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW   ATP-binding; Glycoprotein; Transmembrane; Transport; Phosphorylation.
FT   TRANSMEM   26   42   POTENTIAL.
FT   TRANSMEM   640   656   POTENTIAL.
FT   TRANSMEM   690   706   POTENTIAL.
FT   TRANSMEM   717   733   POTENTIAL.
FT   TRANSMEM   749   765   POTENTIAL.
FT   TRANSMEM   771   787   POTENTIAL.
FT   TRANSMEM   1041   1057   POTENTIAL.
FT   TRANSMEM   1351   1367   POTENTIAL.
FT   TRANSMEM   1661   1677   POTENTIAL.
FT   TRANSMEM   1708   1724   POTENTIAL.
FT   TRANSMEM   1737   1753   POTENTIAL.
FT   TRANSMEM   1775   1791   POTENTIAL.
FT   TRANSMEM   1854   1870   POTENTIAL.
FT   NP_BIND   933   940   ATP (POTENTIAL).
FT   NP_BIND   1946   1953   ATP (POTENTIAL).
FT   MOD_RES   1042   1042   PHOSPHORYLATION (BY PKA) (MAJOR) (BY
FT   SIMILARITY).
FT   MOD_RES   2054   2054   PHOSPHORYLATION (BY PKA) (MAJOR) (BY
FT   SIMILARITY).
FT   CARBOHYD   14   14   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   98   98   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   151   151   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161   161   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   196   196   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   244   244   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   292   292   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   337   337   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   349   349   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   400   400   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   478   478   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   489   489   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   521   521   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   820   820   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD 1144   1144   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD 1294   1294   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD 1453   1453   N-LINKED (GLCNAC. . .) (POTENTIAL).

```
  Query Match            95.1%;   Score 10906;   DB 1;   Length 2261;
  Best Local Similarity  94.8%;   Pred. No. 0;
  Matches 2087;  Conservative   54;  Mismatches   60;  Indels   0;  Gaps   0;
```

```
Qy      1 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDARRLLLYSQKDT 60
          |||||||||||||||||||||||||||||||||||||||||:||||||:|||||||:||
Db     61 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVSRLFSDAQRLLLYSQRDT 120

Qy     61 SMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILH 120
          |:|||  ||||  |:||| :||||||||||||||||| |||||||:|||| :|: :| |
Db    121 SIKDMHKVLRMLRQIKHPNSNLKLQDFLVDNETFSGFLQHNLSLPRSTVDSLLQXNVGLQ 180

Qy    121 KVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLPREKLAAAERVLRSNMDILKPIL 180
          |||||||||||| |||||| ||:||||| ||| |||||:|| |||||||| |||||||::
Db    181 KVFLQGYQLHLASLCNGSKLEEIIQLGDAEVSALCGLPRKKLDAAERVLRYNMDILKPVV 240

Qy    181 RTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMRSWSDMRQEVMFLTNVNSSSSSTQI 240
          |||| |:: ||||  || || ||:| |||||:||| |||||||||||||||||||||||
Db    241 TKLNSTSHLPTQHLAEATTVLLDSLGGLAQELFSTKSWSDMRQEVMFLTNVNSSSSSTQI 300

Qy    241 YQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNGTEEDAETFYDNSTTPYCNDLMK 300
          ||||||||||||||||||||||||||||||||||||||| ||||  :||||||||||||||||
Db    301 YQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNNTEEDVDTFYDNSTTPYCNDLMK 360

Qy    301 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTFQELAVFHDLEGMWEELS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTFQELAVFHDLEGMWEELS 420

Qy    361 PKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTAQDIVAFLAKHPEDVQSSNGS 420
          |:||||||||||||||||:|||||||| |||| || ||||:|||||||||||:||||:||||| |||
Db    421 PQIWTFMENSQEMDLVRTLLDSRGNDQFWEQKLDGLDWTAQDIMAFLAKNPEDVQSPNGS 480

Qy    421 VYTWREAFNETNQAIRTISRFMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTG 480
          |||||||||||||||:||||||||||||||||||| ||| |||||||||||||||||||
Db    481 VYTWREAFNETNQAIQTISRFMECVNLNKLEPIPTEVRLINKSMELLDERKFWAGIVFTG 540

Qy    481 ITPGSIELPHHVKYKIRMDIDNVERTNKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVV 540
          ||| |:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ITPDSVELPHHVKYKIRMDIDNVERTNKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVV 600

Qy    541 EQAIIRVLTGTEKKTGVYMQQMPYPCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIV 600
          |||||||||:|||||||:|||||||||||||||||||||||||||||||||||||||| ||
Db    601 EQAIIRVLTGSEKKTGVYVQQMPYPCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKSIV 660

Qy    601 YEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV 660
          ||||||||||||||||||| |||||||:|||||||||||||||||||||||||||||||||
Db    661 YEKEARLKETMRIMGLDNGILWFSWFVSSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV 720

Qy    661 FLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFAS 720
          |||||:|||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db    721 FLSVFAMVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFSIKIFAS 780

Qy    721 LLSPVAFGFGCEYFALFEEQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTW 780
          |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db    781 LLSPVAFGFGCEYFALFEEQGIGVQWDNLFESPVEEDGFNLTTAVSMMLFDTFLYGVMTW 840

Qy    781 YIEAVFPGQYGIPRPWYFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVS 840
          |||||||||||||||||||||||||||||| ||||||||:|| :||||||||||||:||||
Db    841 YIEAVFPGQYGIPRPWYFPCTKSYWFGEEIDEKSHPGSSQKGVSEICMEEEPTHLRLGVS 900

Qy    841 IQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYIL 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        901 IQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYIL 960

Qy        901 GKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVG 960
              |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GKDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVG 1020

Qy        961 LPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQ 1020
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 LPPSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQ 1080

Qy       1021 GRTIILSTHHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLS 1080
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GRTIILSTHHMDEADILGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLS 1140

Qy       1081 SCRNSSSTVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDI 1140
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 SCRNSSSTVSCLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDI 1200

Qy       1141 GHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAE 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAE 1260

Qy       1201 TSDGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVK 1260
              |||||||||||||||||||||| |||||| |||||||||||||||||||||||||||:|
Db       1261 TSDGTLPARRNRRAFGDKQSCLHPFTEDDAVDPNDSDIDPESRETDLLSGMDGKGSYQLK 1320

Qy       1261 GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQ 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQ 1380

Qy       1321 PWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEWTTAP 1380
              |||||||||||||||| || |||||||||||||||||||||||||||||| |||:||  :|
Db       1381 PWMYNEQYTFVSNDAPEDMGTQELLNALTKDPGFGTRCMEGNPIPDTPCLAGEEDWTISP 1440

Qy       1381 VPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTADILQDL 1440
              |||:|:|||||||||||:|||||||||||||||||||||||||||||||||| |||||:||
Db       1441 VPQSIVDLFQNGNWTMKNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQKTADILQNL 1500

Qy       1441 TGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVNDATKQMKK 1500
              |||||||||||||||||||||||||||||||||||||||||:||||| ||||| ||||
Db       1501 TGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNSQALPPSHEVNDAIKQMKK 1560

Qy       1501 HLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQK 1560
              ||| ||:||||||:||||||| |||:||||||||||||||||||||||||||||||||||
Db       1561 LLKLTKDTSADRFLSSLGRFMAGLDTKNNVKVWFNNKGWHAISSFLNVINNAILRANLQK 1620

Qy       1561 GENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMSFVPASFVVFLIQER 1620
              |||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db       1621 GENPSQYGITAFNHPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFVPASFVVFLIQER 1680

Qy       1621 VSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLAL 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 VSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLAL 1740

Qy       1681 LLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTDNKLNNIN 1740
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||:||
Db       1741 LLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTNNKLNDIN 1800

Qy       1741 DILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSWDLVGRNLFAMAVEG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 DILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSWDLVGRNLFAMAVEG 1860

Qy       1801 VVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGGGQNDILEIKELTKI 1860
              |||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||
Db       1861 VVFFLITVLIQYRFFIRPRPVKAKLPPLNDEDEDVRRERQRILDGGGQNDILEIKELTKI 1920

Qy       1861 YRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLTGDTTVTRGDAFLNRNSILSN 1920
              ||||||||||||||:||||||||||||||||||||:||||||||| |||||||:||||||
Db       1921 YRRKRKPAVDRICIGIPPGECFGLLGVNGAGKSTTFKMLTGDTPVTRGDAFLNKNSILSN 1980

Qy       1921 IHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGEWAIRKLGLVKYGEK 1980
```

```
                  ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      1981  IHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKFGEWAIRKLGLVKYGEK  2040

Qy      1981  YAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVVLT  2040
                 || ||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db      2041  YASNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNCALSIVKEGRSVVLT  2100

Qy      2041  SHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQDFFG  2100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      2101  SHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQEFFG  2160

Qy      2101  LAFPGSVPKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDYSVSQTTLDQVFVNF  2160
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db      2161  LAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDYSVSQTTLDQVFVNF  2220

Qy      2161  AKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV  2201
              |||||||||||||||||||||||||||||||||||||||||
Db      2221  AKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV  2261
```